Contents lists available at ScienceDirect

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

Operational loss modelling for process facilities using multivariate loss functions

Seyed Javad Hashemi, Salim Ahmed, Faisal Khan*

Safety and Risk Engineering Group, Department of Process Engineering, Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL, Canada

ARTICLE INFO

Article history:
Received 16 July 2015
Received in revised form 19 August 2015
Accepted 22 August 2015
Available online 29 August 2015

Keywords:
Process risk assessment
Copula function
Multivariate model
Process safety
Distillation column
Akaike's information criterion (AIC)

ABSTRACT

This paper presents a methodology to develop multivariate loss functions to measure the operational loss of process facilities. The proposed methodology uses loss functions to provide a model for operational loss due to deviation of key process characteristics from their target values. Having estimated the marginal loss functions for each monitored process variable, copula functions are then used to link the univariate margins and develop the multivariate loss function. The maximum likelihood evaluation method is used to estimate the copula parameters. Akaike's information criterion (AIC) is then applied to rank the copula models and choose the best fitting copula. A simulation study is provided to demonstrate the efficiency of the copula estimation procedure. The flexibility of the proposed approach in using any form of the symmetrical and asymmetrical loss functions and the practical application of the methodology are illustrated using a separation column case study.

1. Introduction

Different sources of variations in a process operation, such as feed specifications, wrong settings, control system malfunction and operator error can cause deviation of process variables from the specification limits. The subsequent unprofitable process operation incurs operational loss, which is defined in this work as the loss due to production of sub-quality products and increased energy usage resulting from a deviated process variable. Process facilities possess different characteristics that jointly impact process operational loss. For example, the temperature and differential pressure across a distillation column can be used jointly to monitor the operational loss of the distillation system. Thus, integrated operational loss modelling of process industries requires understanding the joint distribution of all key process characteristics and their correlations.

The loss function approach is widely used to quantify quality loss in the manufacturing industry (Leung and Spiring, 2004; Tahera et al., 2010) by relating a key characteristic of a system (e.g. product composition) to its business performance. More recently, loss functions have been applied to model operational loss for process facilities (Hashemi et al., 2014a). Choosing and estimating a useful form for the marginal loss functions of each process characteristic is often a straightforward task (Hashemi et al., 2014a,b), given that enough loss information from the system is available. For multivariate cases, traditionally, the pairwise dependence between loss functions has been described using traditional families of loss functions. The two most common models occurring in this context are the multivariate quadratic loss function (QLF) (Chan and Ibrahim, 2004; Pignatiello, 1993) and the multivariate inverted normal loss function (INLF) (Drain and Gough, 1996; Spiring, 1993). For instance, Spiring (1993) proposed the following equation for bivariate cases with two parameters for which INLF can be used to represent operational loss:

$$L(Y) = MEL\left[1 - \exp\left\{-\frac{1}{2}(Y - T)^T \Gamma^{-1}(Y - T)\right\}\right] \quad (1)$$

* *Corresponding author*. Tel.: +1 709 864 8939; fax: +1 709 864 3480.
E-mail address: fikhan@mun.ca (F. Khan) .

http://dx.doi.org/10.1016/j.cherd.2015.08.019
0263-8762/

where $\mathbf{Y}$ and $\mathbf{T}$ are 2×1 column vectors of key process characteristics under scrutiny and associated target values, respectively. MEL is the maximum estimated loss and $\mathbf{\Gamma}$ is a 2×2 scaling matrix (shape parameter) relating deviation from target to loss for both parameters. The main limitation of this approach is that the individual behaviour of the marginal loss functions must then be characterized by the same parametric family of loss functions. This restriction has limited their useful application in practical situations. Moreover, other than the QLF and INLF, loss functions usually do not have a convenient multivariate generalization.

According to a review of the existing literature in the area of multivariate loss functions conducted by Hashemi et al. (2014a), it can be concluded that the existing research challenge is to develop a flexible framework to assign appropriate marginal loss functions to key process characteristics irrespective of their dependence structure. Copula models, which provide this flexibility, have begun to make their way into process engineering literature (Hashemi et al., 2015c; Meel and Seider, 2008; Pariyani et al., 2012). Copulas are used to describe the joint distribution of dependent random variables with any marginal distribution. While the theoretical properties of copula functions are now fairly well understood, inference for copula models is, to an extent, still under development (Genest and Favre, 2007).

The contributions of this paper are twofold. First, a new methodology is provided to construct multivariate loss functions using copulas. Second, methodologies are provided to estimate copula parameters and choose the best copula for a specific application. The main objective of this paper is to present the successive steps required to use copulas for modelling the dependent losses and constructing multivariate distributions for specific purposes, including operational loss modelling.

Following the introduction, Section 2 proposes a methodology to develop multivariate loss functions using copula functions. Section 3 reviews the theory of copula functions and Section 4 provides methods to estimate and select copula functions and conduct an uncertainty assessment. A separation column case study is then used in Section 5 to illustrate the practical implementation of copulas, followed by some concluding remarks.

2. Methodology: copula-based multivariate loss functions

It has been shown in earlier studies that the application of the general class of inverted probability loss functions (IPLFs) is a flexible approach to model loss due to process deviations (Hashemi et al., 2014a; Leung and Spiring, 2004). However, the application of IPLFs for systems with multiple key process variables is an existing research challenge due to the restriction in multivariate generalizations. Copula functions are used in this work to overcome this challenge.

Before developing a multivariate loss function, it would be helpful to review the common basis of developing IPLFs. According to Leung and Spiring (2004), let $f(x_i)$ be a probability density function (PDF) possessing a unique maximum at x_i, where x_i represents a key process characteristic (KPC) and $i=1,\ldots,I$ represents I KPCs (e.g. temperature, pressure, composition, etc.). Let $T_i=x_i$ be the value at which the PDF attains its unique maximum, where T_i denotes the target value. Let

Table 1 – Listing of univariate inverted probability loss functions (IPLFs).

Type of loss functions	References	Formulation of loss function[a]
INLF	Spiring (1993)	$L(x,T)=MEL\,\{1-\exp(-(x-T)^2/2\gamma^2)\}$ where $\gamma=\Delta_x/4$
Modified INLF	Sun et al. (1996)	$L(x,T)=\frac{MEL_\Delta}{1-\exp\{-0.5(\Delta_x/\gamma)^2\}}\{1-\exp(-(x-T)^2/2\gamma^2)\}$
IBLF	Leung and Spiring (2002)	$L(x,T)=MEL\{1-D[x(1-x)^{(1-T)/T}]^{(\alpha-1)}\}$ where $D=[T(1-T)^{1-T/T}]^{1-\alpha}$
IGLF	Leung and Spiring (2004)	$L\,(x,T)=MEL\left\{1-(e/T)\,x\exp\left(-x/T\right)^{\alpha-1}\right\}$

[a] MEL_Δ is the estimated maximum loss at distance Δ_x, where Δ_x is the distance from the target to the point where the maximum loss MEL first occurs; x represents the process measurement; T denotes the target value; γ and α are shape parameters.

$\pi(x_i,T_i)=f(x_i)$, $m_i=\sup_{x_i\in X_i}f(x)=f(T)$, and define loss inversion ratio (LIR) as:

$$f_{\mathrm{LIR}}(x_i,T_i)=\pi(x_i,T_i)/m_i. \tag{2}$$

Then, any IPLF takes the form:

$$L(x_i,T_i)=MEL_i\,[1-\pi(x_i,T_i)/m_i] \tag{3}$$

where MEL_i is the maximum estimated loss incurred when the target is not attained. It can be seen from the structure of Eq. (3) that $\pi(x_i,T_i)$ is in the form of a PDF in terms of x_i and T_i, m_i is the maximum of $\pi(x_i,T_i)$; the LIR, $\pi(x_i,T_i)/m_i$, is unitless and has a minimum value of zero when x_i takes on values far from the T_i, and a maximum value of one when x_i is exactly on target, i.e., $0\le\pi(x_i,T_i)/m_i\le1$ (Leung and Spiring, 2004).

Table 1 shows the important IPLFs determined from inversion of Normal, Gamma, and Beta distributions using the method described above. A comparative study of the flexibility of different IPLFs for application in the process industries is provided in Hashemi et al. (2014a).

The same basis as in Eq. (3) is used in this work to develop multivariate loss functions. As shown in Fig. 1, the proposed methodology includes the following steps:

Step 1a: The proposed methodology starts with the identification of key process characteristics (KPCs), x_i, $i=1,\ldots,I$. A KPC is a feature that, if nonconforming, missing, or degraded, may cause unsafe conditions and/or a loss of product quality. For example, operating temperature is the KPC for a polymerization reactor. Different approaches, such as check lists, preliminary hazard analysis (PHA), failure modes and effects analysis (FMEA), fault tree analysis (FTA), hazard and operability study (HAZOP), and master logic diagrams, are often used to identify KPCs (Hashemi et al., 2014b). In this study, it is assumed that the KPCs are known.

Step 1b: The next step is to assign a loss inversion ratio, $LIR_i=\pi(x_i,T_i)/m_i$, to each identified KPC. A least-squares based method to determine the parameters of each LIR is described in Hashemi et al. (2014a).

Step 2: The best copula function and associated parameter(s) should then be selected to represent the dependence structure among identified LIRs.